Probabilistic privacy leakage from Challenge-Response RFID authentication protocols

Seongan Lim, Ikkwon Yie
Inha University
Department of Mathematics
Incheon City
Korea

Abstract: To assure the privacy of RFID Tags, symmetric challenge-response identification protocols have been considered. Due to the low cost requirements for tags, it has been assumed that the tampering of RFID tags is possible. In this paper, we estimate the privacy leakage of challenge-response RFID authentication protocols based on symmetric key using discrete probability under the assumption that tampering the RFID tags are possible.

Key–Words: Challenge-Response Authentication protocol, RFID tag, discrete probability

1 Introduction

Many identification protocols based on symmetric challenge-response have been proposed to assure privacy of RFID tag bearers [6, 3, 4]. Especially in [4], Molnar and Wagner defined a scheme to be 'private' if no adversary is able to distinguish two different tags with different secret keys. We use the term 'private' as Molnar et al. defined.

In general, RFID tags are not assumed to be tamper resistant. Avoine, Dysli, and Oechslin proposed a privacy weakening attack on the private authentication protocol of Molnar et al. by using tampered tags [1]. In this paper, In this paper, we estimate the privacy leakage of challenge-response RFID authentication protocols based on symmetric key using discrete probability under the assumption that tampering the RFID tags are possible. We use only discrete probability and a number of transcripts of executed protocol. Thus our result does not depend on a particular key structure as in Avoine et al.'s attack, or not even on the size of security parameter. In Avoine et al.'s attack, the authors analyze the success probability to identify the target tag using tampered tag and the structure of shared key. In our case, we analyze the success probability of random guess of identifying the target tag using tampered tag.

Suppose that an attacker wants to identify his target tag from a set of n tags given to him. If he can only make random guesses, the success probability would be $1/n$. One may naively think that the success probability of the attacker is still the same for the same quest even if he has tampered tag(s) at hand. We show that this is not so if the attacker's goal is to identify the target tag's transcript from a set of n 'randomly looking' transcripts.

Consider an attacker $A(n,t)$ who can tamper up to t tags and has a list of n tags including his target tag. The attacker obtains n communication transcripts, one for each tag in his list. The objective is to identify the transcript of the target tag. We examine the success probability rigorously, and conclude that the success probability of the correct guess increases by $(t/[(n-1)(n+t)^2])$ from the given transcripts of n tags when the attacker has t tampered tags. This improvement of success probability is possible since all the secrets of the tags, including that of tampered tags are chosen according to a uniform distribution. In RFID system, all the secrets were set when the system established and the attacker cannot determine the secret key of the tampered tags before the attacker tampers a tag. The attacker increment the list of transcripts by adding t transcripts by letting the tampered tags execute protocols with a reader. He can later identify the transcripts belonging to the tampered tags using the secret obtained from tampering. By using this fact, we can set a problem, that is very similar to Monty Hall Problem in discrete mathematics.

Note that Avoine et al. considered the case $n = 2$ and $t = 1$, and the success probability of the attacker is $1/2 + 1/9$ without any knowledge on the structure of the secret key. Without tampered tags, the success probability of the attacker is $1/2$.

We shall consider the attacker $A(n,t)$ who has t tampered tags and show that the success probability of $A(n,t)$ increases to $(1/n) + (t/[(n-1)(n+t)^2])$. We note that the success probability is independent of the security parameter of the underlying challenge-response protocol, it depends only on the number of transcripts and the number of tampered tags given to the attacker.

For a fixed n, the success probability attains its maximum at $t = n$. And smaller n gives higher success probability. Since this success probability is independent of the security parameter of the underlying challenge-response protocol, we conclude that no symmetric challenge-response authentication protocol is private against the attacker $A(n,t)$ if the tags can be tampered and $(t/[(n-1)(n+t)^2])$ is larger than the security margin with respect to random guessing attack of authentication protocol.

2 Probabilistic privacy leakage of challenge-response RFID authentication protocol

Now we consider an attack that is similar to Avoine et al.'s privacy weakening attack on the challenge-response authentication protocol based on symmetric key:

- The attacker has tampered t tags $\tilde{T_i}$, $i = 1, 2, \ldots, t$ and has every secret information of each tampered tags.
- The attacker then choose a target tag T, which is not in the list of the tampered tags.
- The attacker receives a list of n tags $T_{e_1}, ..., T_{e_n}$, among which T is included. He also receives communication transcripts $tr_{e_1}, ..., tr_{e_n}$ from valid authentication protocols, one for each of the tags in the list.
- The attacker's goal is to identify the target tag T's transcript from the communication transcripts. The attacker can use the tampered tags for this purpose.

Note that the attacker's success probability is at least $1/n$. Hence the advantage of the attacker A is

$$Adv_A(k) = |\text{ the success probability of the attack} - 1/n|,$$

where k is security parameter of the challenge-response authentication protocol. In order for an authentication scheme to be private in the sense of Molnar et al., the advantage $Adv_A(k)$ must be negligible in k. If the advantage $Adv_A(k)$ is independent of k and not smaller than 2^{-10}, the upper bound of success probability of on-line guessing attack for authentication protocol [2], then one can say that the authentication protocol is not private, i.e., it has privacy loss.

We assume much weaker power to the attacker than Avoine et al. did. More precisely, our use of the tampered tags are limited to obtain valid transcripts while Avoine et al launches a man-in-the-middle attack to extract information about the hierarchical key structure. In fact, the adversarial action we consider can be applied to any challenge response RFID authentication protocol when tampering tag is allowed.

Let us denote above described attacker as $A(n,t)$ and we index the tampered tags as $T_{-1}, T_{-2}, ..., T_{-t}$ and describe the adversarial action as the following.

Phase 1. Let each tampered tag execute the authentication protocol with a reader to obtain communication transcripts. Together with those n communication transcripts already given, the attacker has $n + t$ 'random' transcripts $tr_{e_1}, ..., tr_{e_{n+t}}$ of $n + t$ tags with $\{e_1, ..., e_{n+t}\} = \{-t, -t + 1, ..., -1, 1, 2, ..., n\}$.

Phase 2 (First guess stage). Choose a transcript at random from the list of $n + t$ transcripts.

Phase 3 (Final selection Stage). Identify those transcripts of t tampered tags using the secrets and ID's of the tampered tags and then discard them.

Case 1: If the transcript chosen in Phase 2 is among the discarded of the tampered tags, then choose a transcript at random from the remaining n transcripts and claim it to be the transcript of the target tag.

Case 2: If the transcript chosen in Phase 2 is not among the discarded of the tampered tags, then discard it also and choose a transcript from the remaining $n - 1$ transcripts and claim it to be the transcript of the target tag.

The success or fail of the attack is determined in Phase 3. The probability of the choice made in Phase 2 falling into Case 1 is $t/(n + t)$ and that of falling into Case 2 is $1 - t/(n + t) = n/(n + t)$. In Case 1, the success probability is $1/n$. The success probability in Case 2 is similar to the Monty Hall problem in the discrete mathematics. The probability for the first choice made in Phase 2 to be the transcript of the target tag is $1/(n + t)$. Thus, in Case 2, the probability for the transcript of the target tag to belong to the remaining list of $n-1$ transcripts is still $1-1/(n+t) = (n+t-1)/(n+t)$. Therefore the success probability in Case 2 is $(n + t - 1)/(n + t) \times 1/(n - 1)$. Thus the overall success probability $P(n,t)$ of the attacker $A(n,t)$ is

$$P(n,t) \quad = \quad \frac{t}{n+t} \cdot \frac{1}{n} + \frac{n}{n+t} \cdot \frac{n+t-1}{n+t} \cdot \frac{1}{n-1}$$

$$\begin{aligned} &= \frac{1}{n+t}\left(\frac{t}{n}+1+\frac{t}{(n-1)(n+t)}\right) \\ &= \frac{1}{n+t}\left(\frac{n+t}{n}+\frac{t}{(n-1)(n+t)}\right) \\ &= \frac{1}{n}+\frac{t}{(n-1)(n+t)^2}\,. \end{aligned}$$

Then the advantage $Adv_{A(n,t)}$ of the attacker $A(n,t)$ is

$$Adv_{A(n,t)} = P(n,t) - \frac{1}{n} = \frac{t}{(n-1)(n+t)^2}.$$

3 Conclusion

In challenge-response protocol, communication transcripts of valid protocols can be regarded as randomly chosen items. Without any aid, the probability of an attacker guessing the correct transcript out of n is $1/n$. By introducing the transcripts of t tampered tags the attacker has advantage in guessing the correct one as much as $t/[(n-1)(n+t)^2]$. Moreover, this advantage is independent of the specific challenge-response protocol or the security parameter.

For fixed n, the advantage $Adv_{A(n,t)} = t/[(n-1)(n+t)^2]$ attains its maximum value $\frac{1}{4(n-1)n}$ when the attacker takes $n = t$. In other words, if an attacker is to identify his target tag among a set of n tags, then he would tamper extra n tags in order to get the best guessing advantage.

For example, in the Library RFID system, it is commonly assumed that the number of tags are 2^{20}. For example, if we have $n = 2^5, t = 2^5$, then the advantage of the attacker $A(2^5, 2^5)$ is

$$Adv_{A(2^5,2^5)} = 1/(2^{17} - 2^12) > 2^{-17}.$$

If the security margin with respect to the random guessing attack of the authentication protocol is required to be smaller than 2^{-12} in the Library RFID authentication protocol, no challenge-response authentication protocol is private against $A(n,1)$.

Another extreme example is the case $n = 2$ as the definition of private scheme of Molnar et al. The attackers $A(2,1)$ and $A(2,2)$ have guessing advantages as much as $1/9$ and $1/8$, respectively. These value $1/9, 1/8$ are much larger than the currently recommended security margin with respect to the guessing probability of authentication protocols. Thus we conclude that no RFID authentication system based on symmetric challenge-response protocol can be private unless tamper resistance is provided if the attacker can set $n = 2$.

Note that the advantage of $A(n,t)$ becomes smaller as n gets larger and it attains is maximum value when $t = n$ for a fixed n. Practically, n may be fairly large while t cannot be assumed to be as large. Once a specific RFID system is given, it may be possible to confine the target tag among a relatively small number n of candidate tags and privacy will emerge as a real issue in such a case. Thus we can conclude that we need to have a tamper resistant RFID tags or need to make n, the size of the transcript sets of distinct tags that the attacker can obtain, as large as the maximal guess advantage $\frac{1}{4(n-1)n}$ be smaller then the security margin of the underlying challenge response protocol.

Acknowledgement

This work was supported by the Korea Research Foundation Grant funded by the Korean Government(MOEHRD)(KRF-2004- R03-10023).

References:

[1] Gildas Avoine, Etienne Dysli, and Philippe Oechslin, *Reducing time complexity in RFID systems*, Selected Areas in Cryptography-SAC '05, Lecture Notes in Computer Science 3897, Springer-Verlag Berlin, 2005, pp.291-306

[2] William E. Burr, Donna F. Dodson, and W. Timothy Polk, *Electronic Authentication Guideline*, NIST Special Publication 800-63 ver. 1.0.2, 2006

[3] Martin Feldhofer, Sandra Dominikus, and Johannes Wolkerstorfer, *Strong authentication for RFID systems using the AES algorithm*, Workshop on Cryptographic Hardware and Embedded Systems-CHES '04, Lecture Notes in Computer Science 3156, Springer-Verlag Berlin, 2004, pp.357-370

[4] David Molnar and David Wagner, *Privacy and security in Library RFID:Issues, practices, and architectures*, Conferences on Computer and Communications Security-CCS '04, ACM, ACM Press, 2004, pp.210-219

[5] Keunwoo Rhee, Jin Kwak, Seungjoo Kim, and Dongho Won, *Challeng-Response based RFID authentication protocol for distributed database environment*, International Conference on Security in Pervasive Computing-SPC '05, Lecture Notes in Computer Science 3450, Springer-Verlag Berlin, 2005, pp.70-84

[6] Stephen Weis, Sanjay Sarma, Ronald Rivest, and Daniel Engles, *Security and privacy aspects of low-cost radio frequency identification systems*, International Conference on Security in Pervasive Computing-SPC ’03, Lecture Notes in Computer Science 2802, Springer-Verlag Berlin, 2003, pp.454-469